EUROPEAN PATENT APPLICATION

(72) Inventor: Pilu, Maurizio
Bristol BS7 ORJ (GB)

(74) Representative: Lawrence, Richard Anthony
Hewlett-Packard Limited,
IP Section,
Building 3,
Filton Road
Stoke Gifford, Bristol BS34 8QZ (GB)

(54) **Curled surface imaging system**

(57) The present invention relates to a system for de-warping images of an applicable surface, including applicable curled surfaces, and in particular of images of curled documents (30). The system (1) includes a processor (25) linked to an image capture means (2) which: captures an image of the surface (30), said image having a warp corresponding to the non-planar surface; generates from the image a first set of data points representing the three-dimensional profile of the non-planar surface relative (30) to a planar reference surface (12); fits to the first set of data points a second set of data points representative of an applicable mesh; and uses the second set of data points to texture-map the image (12) in order to de-warp the image.
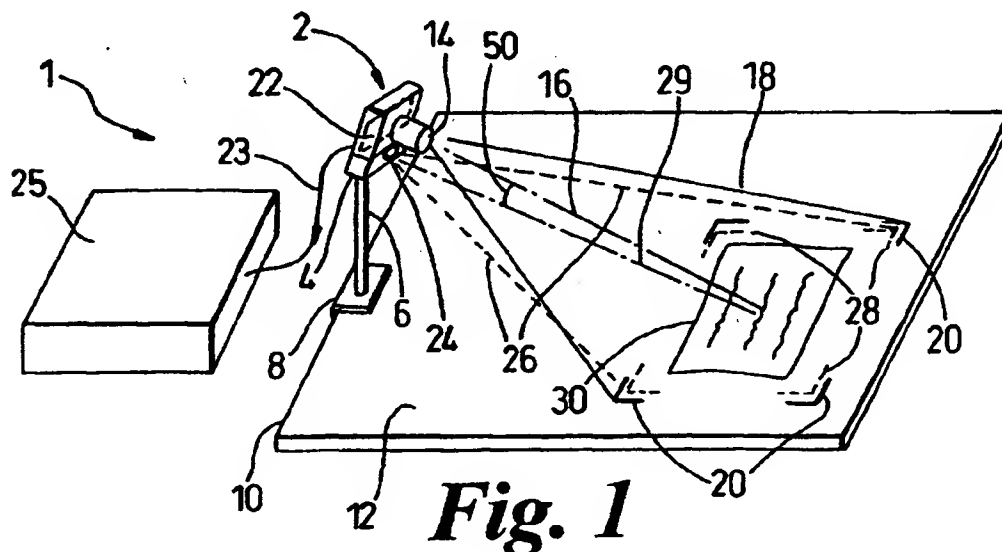
Fig. 1

D scription

[0001]   The present invention relates to a system and a method for de-warping images of an applicable surface, including developable curled surfaces, and in particular of images of curled documents.

[0002]   Images captured of a curled surface will in general exhibit distortions caused by image perspective, skew and compression or elongation caused by an uneven or a curled surface. Standard triangulation techniques can be used to calculate the surface profile from captured images of the surface. For example, a camera can be used with a structured light pattern in a platenless document imaging system to capture the image of a page or of a bound book together with depth information that can be inferred from the light pattern.

[0003]   Desktop flat bed scanners are very common in office imaging applications. Although these are relatively inexpensive and work well, a disadvantage is that these invariably take up a significant amount of desk space, which is always at a premium.

[0004]   Digital camera products are becoming common in many areas of still and motion photography, and as a result are becoming ever less expensive. However such cameras are still used almost exclusively for photography of people or places, and have yet to be adapted for use in office imaging applications. One reason for this is that a document such as a piece of paper or an open book lying face up on a supporting surface is generally not flat, because the document is not held against a transparent platen as in a flatbed scanner.

[0005]   Documents may also not lie at a consistent angle to the camera. In the case of the book, the spine will then be skewed at a variable angle to the optical axis of the camera lens.

[0006]   Therefore, camera-based capture of a document poses the problem of distortion of the captured image due to image perspective, skew and compression or elongation introduced by the uneven surface and page curl of the sheet or bound book.

[0007]   Page curl is one of the biggest problems encountered when capturing a document with a camera. The curled part of the document renders poorly on screen and printers, presents shadows. It is also hard to do stitching and optical character recognition with such a "warped" image.

[0008]   Recovering or "de-warping" page curl is a difficult problem to solve in general. Methods that are known include inferring shapes from shading, from texture, from overall contours of lines of text. These methods have so far proved to be fragile and often require a significant amount of computer processing power.

[0009]   One approach to solve this problem is to use structured light to obtain depth information, such as the distance of the page from a plane at right angles to the optical axis of the camera. Such an approach is disclosed in patent document US 5,760,925, in which a document is supported on an underlying support surface with a cameras mounted on above and to one side of the support surface, and a light stripe projector mounted on an opposite side of the support surface. The light stripe projector projects a pair of light stripes onto the document. The light stripes are parallel for portions of the document the same height above a reference surface, which is taken to be the support surface. The document is oriented so that most of the curl is in the same direction as the light stripes, but because the document may not be flat in a transverse direction, the shape of the document is interpolated linearly between the light stripes.

[0010]   This system can in principle capture an image of the document and correct this for page curl only when there is no curl transverse to the light stripes. Although more parallel light stripes can in principle and at increased cost be added to gain additional curl information in the transverse direction, in practice this places a heavy burden on the available processing power and time available to capture and correct for document curl in a product that is commercially attractive in terms of cost and speed.

[0011]   It is an object of the present invention to address these issues.

[0012]   Accordingly, the invention provides an imaging system for imaging a non-planar applicable surface, the system comprising a processor linked to an image capture means and being capable of: capturing at least one image of the surface, said image having a warp corresponding to the non-planar surface; and of generating therefrom a first set of data points representing the three-dimensional profile of the non-planar surface relative to a planar reference surface, wherein the processor is arranged to fit to the first set of data points a second set of data points representative of an applicable mesh and to use the second set of data points to texture-map the image in order to de-warp the image.

[0013]   Curled paper can be mathematically represented by an applicable or a developable surface that has the property of being isometric with the plane. In practical terms, this means that paper can be uncurled and/or unfolded to a plane without tearing. A set of measured three-dimensional data points representative of a curled applicable surface may consist of scattered and/or noisy data, in which case it is not possible to fit a general averaged surface such as a bicubic spline to the data and unroll, or "texture-map" the surface onto a plane without causing global distortions. The mesh is applicable before texture-mapping and therefore recovers at least to some extent the original applicable surface profile of the non-planar surface. In the case of a document imaging system, this p rmits the image of the document to be de-warped.

[0014]   Also according to the invention, there is provided a method of imaging a non-planar applicable surface using an imaging system comprising a processor linked to an image capture means, comprising the steps of:

i) capturing at least one image of the surface said image having a warp corresponding to the non-planar surface;

ii) generating from the image a first set of data points representing the three-dimensional profile of the non-planar surface relative to a planar reference surface;

iii) fitting to the first set of data points a second set of data points representative of an applicable mesh; and

iv) using the second set of data points to texture-map the image in order to de-warp the image.

[0015] In a preferred embodiment of the invention, in step iii) the mesh is distorted as the second set of data points is fit to the first set of data points to the extent that the mesh is no longer applicable, following which the distorted mesh is relaxed to an applicable state. This permits the mesh to average out deviations from an applicable state in the first set of data points, which can result in a better recovery of the original applicable surface from the scattered and/or noisy data.

[0016] Preferably, prior to step iii) an "initial" surface is fit to the first set of data points, and in step iii) the mesh is fit to the initial surface. For example, the initial surface may be a bicubic spline surface. This in general will not be an applicable surface, in which case the mesh will become distorted as it is fit to the initial surface.

[0017] However, in a preferred embodiment, the distortion of the mesh takes place in two stages so that the second set of data points may be better fit to the first set of data points. Here, the initial surface may be an applicable surface such as a plane, fit in a least squares routine to the first set of data points. The mesh is then not distorted in a first stage when it is fit to the initial surface, but rather in a second stage in which after fitting of the mesh to the initial surface at least some of the second set of data points are moved closer to corresponding ones of the first set of data points during which the mesh is distorted.

[0018] If some data points in the second set of data points do not correspond closely enough to any of the data points in the first set of data points, then these data points need not be fit the first set of data points.

[0019] In a preferred embodiment of the invention, the relaxation of the mesh takes place in an iterative process in which the second set of data points is adjusted incrementally until distances between points in the second set of data points are equalized.

[0020] The image capture means may project a structured light pattern that forms separated light stripes across the non-planar applicable surface, the first set of data points being generated from the light stripes. Then, step ii) may include the steps of:

a) creating a difference image by taking a difference between an image captured with the stripes and an image captured without the stripes;

b) thresholding the difference image to discard portions below a threshold;

c) counting detected stripes across the difference image in order to identify individual stripes

d) triangulating the image of the non-planar surface at points corresponding with identified stripes to generate the first set of data points.

[0021] The invention will now be described in further detail by way of example only, and with reference to the accompanying drawings, in which:

Figure 1 is a schematic view in perspective of a document imaging system according to the invention, with a camera having a detector array mounted together with a light stripe projector on a post overlooking a document to be imaged;

Figure 2 is a view of an open book as imaged by the camera;

Figure 3 is a view of the open book with a light stripe pattern projected onto the book;

Figure 4 is a schematic view in perspective showing a structured light pattern produced by the light stripe projector showing diverging sheets of light that bow concavely inwards toward a central planar light sheet;

Figure 5 is a plot of light stripes formed by an intersection of the structured light pattern of Figure 3 with a plane transverse to the sheets;

Figure 6 is a plot of a polynomial fitted through five points taken from one of the light stripes of Figure 5;

Figure 7 is a plot of a parametric light sheet cone constructed from polynomials such as those of Figure 6;

Figure 8 is a flow chart depicting a calibration process for the roto-translation between the camera and the light stripe projector;

Figures 9A and 9B show the error in detection of the profile of a planar surface using the document imaging system when an initial estimate of the roto-translation is used;

Figures 10A and 10B show the error in detection of the profile of a planar surface after calibration of the roto-translation according to the flow chart of Figure 8;

Figure 11A and 11B show respectively the errors in the detection of a planar surface without and with radial distortion correction;

Figure 12 shows the results of a stripe detection and labelling process;

Figure 13 shows a light sheet and terminology used in a triangulation process to compute the intersection of the light sheet and the light stripe;

Figure 14 shows a set of measured data points of an open book generated with the document scanning system;

Figure 15 shows the profile of a surface fitted with a bicubic spline to the measured data points according to a prior art method of dewarp curl;

Figure 16 show the prior art results of unrolling the bicubic spline onto a plane;

Figure 17 is a schematic diagram of a prior art orthoimage method to dewarp curl when applied to the fitted surface;

Figure 18 is a schematic diagram of a triangular mesh used to approximate an applicable surface;

Figure 19 is a schematic diagram of a way of estimating from the measured data points the approximate extent when dewarped of a curled document;

Figures 20A, 20B and 20C show in a two-dimensional analogy how the mesh is initially fit to the measured data points in a process which stretches the mesh so that it is no longer applicable;

Figure 21 is a schematic diagram showing by analogy with a spring mesh how the mesh is relaxed back to an applicable state in which it is optimally fit to the measured data points; and

Figure 22 shows a process of texture-mapping the relaxed mesh to de-warp the curled image of the document.

[0022]   Figure 1 shows a document imaging system 1 that has an electronic camera 2 a lower portion of which houses a light stripe projector 4 manufactured by Lasiris, Inc. of St. Laurent, Quebec, Canada, as model number 515L. The camera 2 is mounted atop a support 6 that is clamped 8 to and rises above an edge 10 of a work surface 12. The camera 2 has a main lens 14 with an optical axis 16 that is directed across and down upon the work surface 12. The lens 14 has a field of view 18 that images an area 20 of the work surface 12 onto a two-dimensional CCD detector array 22 within the camera 2.

[0023]   The detector array is connected 23 to a processor unit 25, which may, for example, be a personal computer with an expansion card for controlling the camera 2, light stripe projector 4, and for receiving and processing data received from the detector array 22.

[0024]   Ideally, the area 20 is at least of A4 document size. Similarly, the light stripe projector 4 has a projecting lens 24 that projects a structured light pattern 26 onto a work surface area 28 that is roughly coincident with the imaged area 20. The structured light pattern will be described in more detail below, but extends around a central axis 29 that is roughly coincident on the work surface 12 with the camera lens axis 16. The spread of the structured light pattern is sufficient to cover an A4-size area at about 300 mm distance.

[0025]   A document 30 has been placed within the area 20,28 defined by the camera lens 14 and structured light

pattern 26. The document is supported by the work surface 12 in a generally horizontal orientation, but is slightly curled. An image captured by the detector array 22 will therefore have perspective foreshortening owing to the oblique angle between the camera optical axis 16 and the document 30, and well as warp distortion due to the document curl.

[0026] Such warp distortion can be seen in Figure 2, which illustrates an image 31 of an open book 32 as formed on the detector array 22 by the camera 2. The amount of warp distortion is greatest near the spine 34.

[0027] Figure 3 shows an image 33 of the open book 32 when the structured light pattern 26 is projected towards the book 32 to produce fifteen separated light stripes 35 over the book 32. The book is oriented so that the light stripes 35 are transverse to the spine 34.

[0028] A difference can then be taken between the image 33 with the light stripe pattern 35 and the same image 31 without the light stripe pattern 35, in order to detect the light stripes.

[0029] As can be appreciated from Figures 1 and 3, the advantage of having the light stripe projector mounted together with and below the camera is that the furthest stripe 36 will always be in view of the camera, even if the stripe is projected beyond the further edge of a book.

[0030] Figure 4 shows how the structured light pattern 26 is produced. A fixed 7 mW laser 38 projects a beam of light 40 to a first optical diffractive element 42 that diffracts the beam 40 into a vertically oriented strip 44. The vertical stripe 44 is then diffracted by a second optical diffractive element 46 into the structured light pattern 26 consisting of fifteen diverging, separate and non-intersecting sheets of light 48.

[0031] The structured light pattern 26 is projected onto the document 30 with the projection axis 29 offset at an angle 50 to permit triangulation of the light stripes 35 to characterise document curl.

[0032] These diffractive elements 42,44 produce a set of seven conical light sheets 51,53 either side of a central planar light sheet 52. The central planar light sheet 52 contains a median ray 49, which also lies on the light stripe projector axis 29.

[0033] Each set of light sheets 51,53 bows concavely inwards towards the central planar light sheet 52, with the result that the divergence between adjacent light sheets is a minimum at the middle of the light sheets 48. The sheets are symmetric about a plane that is transverse to the planar sheet and which comprises a median ray of the planar sheet 52.

[0034] As shown in Figure 5, the conic light sheets 51,53 will in general generate curved, non-parallel light stripes on the document, with a concentration of light stripes along a line centrally transverse to the light stripes 35. The concentration of stripes corresponds with the minimum divergence between adjacent light sheets. In Figure 3, this concentration is about the book spine 34. In this example, the concentration of light stripes about the spine 34 will provide enhanced curl information in the region of greatest document curl.

[0035] Triangulation of conic light sheets is, however, a non-trivial problem. For this reason a closed-form solution to this triangulation problem is described below that can be applied in general with this kind of structured light to characterise document curl. The closed form of the triangulation also allows the use of a standard optimisation method to perform an initial calibration of the camera 2 and light stripe projector 4 system.

[0036] Although the use of a multiple line structured light pattern has advantages in terms of cost, the time needed to capture an image, and mechanical complexity over traditional laser scanning methods in the sense that there are no moving parts such as galvanometers or stepper motors, there is a drawback in that the three-dimensional resolution is less, being limited to the number of lines in one direction. This drawback is partly mitigated by the concentration of lines in the region of greatest curl and, as will be explained below, by the use of novel methods to characterise and de-warp image curl.

[0037] Triangulation works as follows. First, light stripes 48 are projected onto an object which is viewed by and projected onto the camera image plane at the detector array 22. Let us suppose for the moment that the laser projects just a single light sheet of a known shape and defined by a corresponding known equation in the camera reference system, which when cast onto an object and imaged by the camera produces a single curve (or stripe) on the image plane. A given point of the stripe defines a line in space going from the camera optical axis through the image. The intersection between the light sheet and this line defines a three-dimensional point in the camera reference system that is on the surface of the object. By repeating the procedure for each stripe point, we can effectively recover all the object's points that lie on the curve defined by the intersection of the projected light sheet and the object surface.

[0038] In this invention we do not have a single light sheet but rather a set of them slightly displaced in order to cover a larger portion of the object, and thus obtain a three-dimensional snapshot of it. In the single light sheet case we knew that any image stripe point corresponded to the projection of a 3D point of a known stripe and this is why it is possible to do triangulation unambiguously. Conversely, in the multiple light sheet case we do not actually know which particular light sheet generated that projection and so some sort of stripe labelling or identification is necessary to determine which image stripe was generated by a particular light sheet.

[0039] The camera 2 and light stripe generator 4 system is initially calibrated by measuring a reference surface, which for convenience may be a support surface 12. The vertical displacement between the undistorted projection on the reference surface and the distorted projection on the curled document is a unique function of depth or height of

the curled document relative to the reference surface.

[0040] The diffractive light stripe proj ctor 4 produces a structured light pattern that with stripes 48 each of which that has a periodic intensity variation along its length. To a first approximation, the peaks in light intensity of the structured light pattern therefore occur at points, which on a spherical surface centered on the light stripe projector can be represented by the following equations:

$$x_{i,j} = i\lambda \frac{D}{\Lambda_1 N_{i,j}}$$

and

$$y_{i,j} = j\lambda \frac{D}{\Lambda_2 N_{i,j}}$$

where

$$N_{i,j} = \left[1 - \left(i\frac{\lambda}{\Lambda_1} \right)^2 - \left(j\frac{\lambda}{\Lambda_2} \right)^2 \right]$$

(1)

and where (x,y) = (0,0) is on the projection axis 29, D is the distance from the light stripe projector 4, λ is the laser wavelength, Λ_1 is the period of the grating for diffractive element 42 and Λ_2 is the period of the grating for diffractive element 46.

[0041] Figure 5 shows fifteen light stripes formed by the intersection of a plane spaced 0.5 m from the light stripe projector and at right angles to the light sheet projection axis 29. The central planar light sheet 52 produces a straight light stripe 54, and light stripes 55,57 on either side of the central light stripe 54 bend inwards towards the central light stripe 54. The light stripes are therefore concentrated along a central line 56 transverse to the central stripe 54.

[0042] In order to perform triangulation in closed-form on the projected light stripe pattern, it is necessary to express this pattern and hence each light sheet 48 in a mathematical form. Therefore, five points 58 which correspond to subsidiary maxima along each light stripe 50 are used, as shown in Figure 6, to generate a second-order polynomial of the projected stripe on the orthogonal plane at a given distance. Although the stripe is, strictly speaking, not quadratic, we have noticed that the deviation from the data is less than 0.01% when the polynomial is of second order. The equation of the polynomial Γ can be expressed in parametric form as:

$$\Gamma_N : \begin{cases} x_{\Gamma,N} = u \\ y_{\Gamma,N} = b_N + a_N u^2 \\ z_{\Gamma,N} = c_N \end{cases}$$

where the index N stands for the stripe number and u is a free parameter illustrated graphically in Figure 7. From this we can construct a cone 70 centered on the projection axis 29, by letting v be a parameter sweeping the cone length. The cone 70 is expressed as:

$$\Sigma_N : \begin{cases} x_{\delta,N} = v x_{\Gamma,N} = vu \\ y_{\delta,N} = v y_{\Gamma,N} = v(b_N + a_N u^2) \\ x_{\delta,N} = v z_{\Gamma,N} = v c_N \end{cases}$$

(2)

[0043] As explained below, of particular interest is the algebraic form of each cone of light, which is obtained by elimination:

$$\Sigma_N: \; yz\text{-}z^2 \frac{b_N}{c_N} \text{-} a_N c_N x^2 = 0$$

[0044] In order to perform triangulation in closed-form, it is necessary also to know the relative orientation of the camera lens axis 16 and the light stripe projector axis 29, referred to herein as the roto-translation R_{OL} between the camera 2 and the light stripe projector 4.

[0045] The intrinsic camera model employed in this invention is described by a conventional set of five parameters, which are the focal length f, the number of pixels per meter in the horizontal and vertical direction α_x and αy, the "piercing point" $(x_0 \; y_0)$ (assumed to be at the image centre) plus the radial distortion parameter K.

[0046] The calibration of the camera parameters f, α_x and αy, the "piercing point" (x_0, y_0) can be estimated with a method described by Tsai, R. Y., IEEE Transactions on Robotics and Automation, No. 4 pp. 323-344, 1987.

[0047] The estimation of the roto-translation R_{OL} is accomplished by a custom method based on an optimization process starting with the capture of sparse three-dimensional data of a planar object. An initial rough estimate of R_{OL} is then determined. Following this, an iterative process shown in Figure 8 is used to adjust six parameters representative of R_{OL} (three Euler rotations angles and three translations) until triangulated data points become effectively planar. Minimization of errors is carried out by an implementation of the Levenberg-Marquard method.

[0048] Figures 9A and 9B show two graphs that illustrate the errors in measuring a planar surface using the initial rough estimate of R_{OL}. Figures 10A and 10B show two similar graphs using the final estimated R_{OL} after the optimisation process of Figure 8. These graphs show that the reduction in the standard deviation of the error in the measurement of the plane is reduced from 20 mm to less than 1 mm. The residual error is due to measurement noise.

[0049] Correction of radial distortion is generally neglected in the field of document imaging. However, it has been discovered that this correction is important in order to obtain sufficiently accurate results. The mapping from distorted to undistorted co-ordinates is:

$$\begin{cases} \bar{x} = x + xK_1 d^2 \\ \bar{y} = y + yK_1 d^2 \end{cases} \quad d = \sqrt{(x-c_x)^2 + (y-c_y)^2}$$

[0050] For simplicity of presentation, these new coordinates will in the following description be treated as the actual image coordinates, although one has to bear in mind that these are corrected coordinates derived from the above mapping.

[0051] The camera 2 used in the present example has a radial distortion parameter $K = 0.004\text{pixels/mm}^2$. Figure 11A shows how, even when the R_{OL} has been calibrated, if the radial distortion is not accounted for the error becomes very large. Once this distortion is allowed for, the distortion is as shown in Figure 11B.

[0052] Because there is more than one light stripe, it is necessary to identify each light detected stripe before triangulation is performed. There are two distinct parts in this process, the first one being stripe detection and the second one stripe labelling.

[0053] The three-dimensional document image capture can be done by briefly flashing the laser pattern and synchronously detecting with the detector array 22 the document image including the light stripe pattern, as shown in Figure 3. Either before or after this, the document is imaged without the light stripe pattern, as shown in Figure 2. There will then be two overlapping images, one with the pattern and one without and thus it is straightforward to use image differencing to make the stripes stand out.

[0054] However, the intensity value across stripes will in general be uneven, for example owing to subsidiary peaks as in equation (1) above, or because of uneven ambient illumination or paper reflectance. Therefore, the image of the lights is processed. Given the prevalently horizontal lines, the first step is to use a one-dimensional Laplacian operator (second derivative) applied only on the y (vertical) direction. The application of this operator gives the centre of the stripe a high negative value. This value can then be thresholded to obtain a binary image. The process is robust and fast but the use of a single threshold may inevitably cause some gaps in the continuity of the detected stripes, as shown in Figure 12.

[0055] The method used to identify light stripes is as follows. First, the stripes are thinned down to one pixel thickness and connected pixels are joined together into a string. Next, strings that are too short are removed from the classification

as deemed noise. The result is a data set of pixels as shown in Figure 12, where string segments 80 are interspersed with gaps 82.

[0056] Then for each string, a heuristic "strength" measure is computed as:

$$S=0.5*Length + 0.5*Abs(Avg(Top\ 30\%\ of\ Laplacian\ value))$$

[0057] This is an equally weighted sum of the length and the average of the top third of the absolute value of the Laplacian values. We do not average all the values of the Laplacian along a string because the stripe intensity is not uniformly distributed and some faint sections might adversely affect the average.

[0058] Next, for each column of pixels and starting from the top row of pixels of the image, we assign successive, increasing label numbers to, and only to, the strongest stripe points in the sense above. The numbering of stripes stops at the maximum number expected, here fifteen. Finally, for each string we assign a label equal to the most popular label assigned to all the points of that string. Figure 12 shows the labelling result in which all stripes are correctly identified.

[0059] This approach, which is essentially a voting method, is very robust in general situations and can smoothly cope with gaps. It is also relatively fast to run with inexpensive hardware. It has the advantage that the light stripes are individually indistinguishable, with individual stripes being identified by counting the series of stripes.

[0060] There are situations in which the method would fail to label stripes properly, namely when the assumptions onto which it is based are not meet. For instance if a stripe is completely or largely invisible or occluded by, for example, a thick book edge, the label assignment will be meaningless. Therefore, in an alternative embodiment not illustrated in the drawings, the light stripes are made individually distinguishable by spatial modulation.

[0061] Having said so, these are situation that should not occur in practice when the light stripe projector is properly arranged with respect to the camera, for example being mounted on the same side of the document as the camera and below the level of the camera.

[0062] The proposed approaches to identifying stripes are quick and simple compared with other approaches in which the stripes are temporally modulated or otherwise made individually distinguishable, for example by colour coding.

[0063] Three-dimensional data points can then be obtained via triangulation, which as illustrated in Figure 13 consists of finding the intersection between the sheet of light 48 and an optic ray going 84 through a given point 86 on the projected stripe 88 and a corresponding point 90 on the detected image 92 in the detector plane 94.

[0064] Referring to Figure 13, let $^oP = (X,Y,Z)$ be a three-dimensional point in the camera reference system O, $^op = (x,y)$ a stripe point in the image plane, $^L\Sigma$ the conical surface representing the conical light sheet 48 in the light stripe projector reference system L, and R_{OL} the transformation between the two reference systems expressed by the four-by-four matrix:

$$R_{OL} = \begin{bmatrix} n_x & o_x & a_x & p_x \\ n_y & o_y & a_y & p_y \\ n_z & o_z & a_z & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

[0065] The triangulation problem is to finding the intersection between a generic elliptic cone and a line in space. First we transform the cone into the reference system of the camera via R_{OL} by expressing a cone point in terms of a point in the O reference system transformed into L:

$$^o\Sigma : \left\{ \left\{ ^L\Sigma : {}^Ly^Lz - {}^Lz^2 \frac{b_N}{c_N} - a_N c_N {}^Lx^2 = 0 \right\} \leftarrow \left\{ \begin{matrix} {}^Lx=n_x^*x+o_x^*y+a_x^*z+p_x \\ {}^Ly=n_y^*x+o_y^*y+a_y^*z+p_y \\ {}^Lz=n_z^*x+o_z^*y+a_z^*z+p_z \end{matrix} \right\} \right\}$$

[0066] The parametric form of the optic ray is:

$$^{\circ}x = tx_i$$
$$^{\circ}y = ty_i$$
$$^{\circ}z = tf$$

5 where f is the focal length and the x and y are expressed in image coordinates. Then we write down a system that expresses the intersection between this cone and the optic ray:

10

$$\begin{cases} ^{\circ}\Sigma : f(^{\circ}x, ^{\circ}y, ^{\circ}z) = 0 \\ ^{\circ}x = tx_i \\ ^{\circ}y = ty_i \\ ^{\circ}z = tf \end{cases}$$

15

[0067] By simple substitution, we arrive at a second order equation in the parameter of the optic ray t:

$$At^2 + Bt + C = 0 \rightarrow [t_1, t_2]$$

20

whose solutions t_1 and t_2 represent the two intersections of the ray with the cone. This equation can be solved analytically and the rather knotty solution has been found but is omitted here for clarity.

[0068] We are interested in only one of the above-mentioned intersections which turns out to be, because of the way we constructed the cone, the one corresponding to the smallest parameter u spanning on the half cone closer to the

25 Z axis of the reference system L.

[0069] Hence, we transform both solutions back to the light stripe projector reference system L:

30

$$^{L}P_{\lambda} = \begin{bmatrix} ^{L}X_{\lambda}/w \\ ^{L}Y_{\lambda}/w \\ ^{L}Z_{\lambda}/w \\ 1 \end{bmatrix} = R_{OL}^{-1} \begin{bmatrix} t_{\lambda}x_i \\ t_{\lambda}y_i \\ t_{\lambda}f \\ 1 \end{bmatrix}$$

35

and use the 2nd of Equations 2 to recover the two corresponding u's, that is:

40

$$\bar{t} = \left\{ t_{\lambda} \text{ such that } \lambda = \arg\min_{\lambda}\left(u_{\lambda} = \frac{^{L}X_{\lambda}c_N}{^{L}Z_{\lambda}} \right), \ \lambda = 1,2 \right\}$$

45 [0070] Finally, the three-dimensional intersection point is given by:

$$X = \bar{t}x_i$$

50

$$Y = \bar{t}y_i$$

$$Z = \bar{t}f$$

55

[0071] We have now found the coordinates of the point in space that belongs to the intersection of the light sheet with the object and whose projection is a particular stripe pixel in the image.

[0072] This process has to be repeated for each pixel (and optionally at sub-pixel resolution) of each of the fifteen

stripes. The triangulation process is very fast but if necessary it would be possible to sub-sample along each line. The result is a "cloud" of three-dimensional data points such as that shown in Figure 14.

[0073]   Now we have a cloud of three-dimensional points 95 representing the paper surface. The problem is how to use these data points 95 to undo or "de-warp" the curl distortion.

[0074]   It is in general difficult to de-warp an image of a curled document. The main problem is that paper is not a generic surface, but a "developable" surface, that is one that can be unfolded without tearing or stretching onto a plane. The Gaussian curvature K of a developable surface S(u,v) is zero everywhere, i.e. K(u,v) = 0. An applicable surface is a more general case of a developable surface, in which there are sharp creases.

[0075]   The conventional surface reconstruction approach of fitting and regularizing a surface, possibly with some discontinuities, does not apply to our problem since we not only need to reconstruct, but we also have to unfold this surface onto a plane. This is not possible if the reconstructed surface is not applicable in the first place. Hence, it is necessary to constrain the fitted surface to be applicable, that is, with zero Gaussian curvature everywhere, which is a not trivial operation.

[0076]   Figures 15 and 16 illustrate why a simple approach will in general not work. The three-dimensional data of Figure 14 has been smoothed and a bi-cubic spline surface 96 has been fitted. In the ideal case where data is noiseless and the light stripe projector and camera system is perfectly calibrated, a fitted surface should also be applicable, but in reality the surface we obtain is clearly not so. For example, see the little bumps in some places 98.
If we now uncurl the page, we have to texture-map patches from the original image onto patches of a plane, a mapping computed by integration of finite differences in the meshed surface 96 as shown in Figure 16.

[0077]   However, by definition, a non-applicable surface can only be unrolled onto a plane by either tearing or stretching which causes unnatural distortions in the unrolled document 100. This is due to the integrative nature of unrolling a surface where locally small errors tend to build up and lead to unsightly distortions. Figure 16 shows the distortion in the texture, which his caused by the irregularities in the farther side of the reconstructed planar mesh in Figure 15.

[0078]   So the problem of unrolling the page can be restated as a problem of fitting an applicable surface onto noisy data.

[0079]   A second problem is that the light stripes do not cover the entire page or there might be gaps right near the edges of the page/book. In this case we do not have three-dimensional data so we would not know how to unroll these regions.

[0080]   Briefly, the method used with the present invention uses a finite element model represented as a triangular mesh that simultaneously fits to the data and constrains the data to be isometric with a plane (i.e. applicable) by a relaxation process.

[0081]   First consider the problem in two-dimensions as illustrated in Figure 17. Here there is a first set of points 102 representing noisy measurements of a curve along its length. A second set of points 104 can then be fit with a least squares fit to the first set 102. A connected piecewise linear curve 106 can be constructed going through the second set of points 104. The second set of points 104 can always be "undone" to a line 108, as the linear curve is isometric to a line. This property explains why many methods that seek to undo the page curl of books use a one-dimensional model of the curl and produce good results when the document curl is essentially cylindrical. However, in a general three-dimensional case it is not possible simply to unfold a two-dimensional set of data points representing a curled document owing to noise or other inaccuracies in the data points.

[0082]   There is a old technique used in cartography called orthoimage projection which is described here for its relevance to this work. The method essentially does not correct for page curl but simply projects the texture orthographically onto a plane. This method, albeit simple and not prone to local distortions, is fundamentally flawed, because it does not unfold the document but rather just "pushes it down.

[0083]   A surface is called a developable surface when its Gaussian curvature vanishes at every point. Developable surfaces can be flattened onto a plane without stretching or tearing. Similarly, a developable surface is one that is obtained by bending a plane, where with bending we mean a transformation that preserves arc length.

[0084]   Note that not all ruled surfaces are developable. Developable surfaces are a special subclass of ruled surfaces, that is surfaces that are generated by a straight line moving in space.

[0085]   A developable surface is a subclass of the more general applicable surfaces. In practice, applicable surfaces have the same properties as developable surfaces, but may have creases.

[0086]   The analytic form of a developable surface is a parametric equation of a ruled surface with the constraint of the tangent plane being the same along each ruling. This definition is *per se* impractical and is mainly suitable for interactive modelling or display.

[0087]   A Finite Element Model (FEM) can be used to represent an applicable or a developable surface, for example a mesh 110 such as that shown in Figure 18 having triangular tiles 111. Such a mesh can be deformed 112 to approximate 114 a developable or an applicable surface. When the mesh 100 is deformed, the til s 111 remain unchanged.

[0088]   A developable or applicable surface can be modelled with a triangular mesh by assuming that the lengths of mesh edges 116 between mesh nodes 117 keeps constant as the mesh is deformed. Of course, making the mesh finer

can make any approximation error arbitrarily small. It has to be noted, however, that it is in general not possible to split triangles and refine the mesh locally to reduce the error in, say, regions of high curvature once the mesh has started deforming.

[0089] Creas s too can be modelled with such a deformable wire-frame mesh. In fact, by increasing the resolution of the mesh it is possible to model more and more accurately any applicable surface.

[0090] The document curl characterisation process described above will in general produce noisy, sparse data as shown in Figure 14. The extent of the surface may not be known. Figure 19 shows one way to estimate the extent of the surface. A convex hull or rectangle 118 (or equivalently a square) enclosing all the data points 95 is projected onto the support plane 12. A rectangle 124 can then be deduced from extreme points of the projected lines 126,128. In Figure 19, a B-spline fits the data 95 and estimates its extent by integration along some chosen curves 120,122.

[0091] Alternatively, the document scanning system may permit a user to select the size of the document. Alternatively, the extent could be determined straight away from the image only using the known system geometry, that is, stand position with respect to camera 2 and camera calibration parameters. This latter approach would also help overcome problems of mismatch between the structured light pattern and the document dimensions.

[0092] Such a mismatch could occur if part of the part of the three-dimensional data 95 does not belong to the same developable surface. This might be the case if the data is of a thick book or a small curled document and the structured light pattern is bigger than the document region. In this case there needs to be a way to tell what belongs to the document and what is does not. This could be done by analyzing a generic surface fitted to the data with a search for steep curvature or depth changes. Data points outside such sudden changes could then be discarded. Another way would be to allow for the mesh to break during the "relaxation" process described below wherever the process does not converge.

[0093] Once the extent of the surface 118 and the corresponding planar projection 124 are known, the mesh 110 is fit to the noisy set of data points 95. The process can be understood with reference to Figures 20A, 20B and 20C, which show for clarity an "initialization" process in a two-dimensional analogy.

[0094] First an "initial" surface, here a plane 130, is fit with a least squares deviation through the noisy three-dimensional set of data points 95. Then the planar mesh 110 is rotated and translated so as to coincide with this plane 130 and the estimated extent 124 of the surface 118. Then each mesh node 117 is vertically translated 132 at right angles to the least squares fit plane 130 towards the closest data point 95. If there is no data point 95 within a certain radius, here taken to be one-third the distance to the nearest neighbouring node 117, then the node is left in place, as is the case in Figure 20C for one node 134. The result is a distorted mesh 133.

[0095] At this stage, the mesh 133 is no longer applicable, that is the isometry with the plane 130 we started with is lost. However, albeit coarsely, the mesh 133 does now approximate the surface 118. The next stage is to adjust the mesh 113 so that it is again applicable, and this is done in a "relaxation" process.

[0096] Let us first define the terminology to be used. Let

$$\mathbf{x}_i = \begin{bmatrix} x_i & y_i & z_i \end{bmatrix}^T$$

be a mesh node defined as a vector of corrdinates in a Cartesian system, and let

$$X = \left\{ \mathbf{x}_1, \ldots \ldots \mathbf{x}_N \right\}$$

be the set of nodes of the mesh. Also let e_{ij}, be an edge of the mesh joining two nodes x_i and x_j and let

$$E = \left\{ e_{i_1 j_1}, \ldots \ldots e_{i_M j_M} \right\}$$

be the set of all the edges of the mesh. The mesh can then be represented by $M = \{X,E\}$. Let us also define a neighbourhood of a node x_i as the set of nodes

$$N_i = \left\{ x_j : e_{ij} \in E \right\}.$$

[0097] We shall indicate with $d_{ij} = \sqrt{(x_i - x_j)^T (x_i - x_j)}$ the Euclidean distance between two nodes and with $\hat{d}_{ij}$ the ref-

erence distance that the mesh had in its original flat state.

[0098] In order to transform the mesh to an applicable state while still approximating the data, an optimization method is used to minimize th deviation of the mesh from the ideal applicable state. Figure 21 illustrates by way of a mechanical analogy a mesh 140 of springs 142. In a relaxed state, the mesh 140 has relaxed springs 142 of extension d_{ij} connected to each other at nodes 144. This reticular structure 140 will therefore be at a minimum energy in a stable state when all the springs have extension d_{ij} and when this happens the mesh 140 is isometric with the plane.

[0099] Hence, the problem is equivalent to that of minimizing the total elastic energy of the system:

$$U = \frac{1}{2} K \sum_{i=1..N} \sum_{j \in N_i} \left[\hat{d}_{ij} - d_{ij} \right]^2$$

[0100] This is done using the well-known gradient descent methoc that iteratively adjusts the position of the nodes until the final, lowest energy is reached. Note that the elastic constant K can be ignored during the minimization process.

[0101] Node co-ordinates are updated according to the followinc rule:

$$x_i^{t+1} = x_i^t + w_{xi} \frac{\partial U}{\partial x_i}$$

$$y_i^{t+1} = y_i^t + w_{yi} \frac{\partial U}{\partial y_i}$$

where w is a factor that will be discussed later.

[0102] Convergence is reached when all the displacements fall below a set treshold.

[0103] The derivatives are straightforward to compute and are:

$$\frac{\partial U}{\partial x_i} = \sum_{j-l_i} (\hat{d}_{ij} - d_{ij}) \frac{(x_i - x_j)}{d_{ij}} \qquad \frac{\partial U}{\partial y_i} = \sum_{j-l_i} (\hat{d}_{ij} - d_{ij}) \frac{(y_i - y_j)}{d_{ij}}$$

[0104] Note that these derivatives could be also rewritten as the resultant of the forces exerted to each one of the nodes 144 by all the springs 142 connected each particular node.

[0105] Regarding the convergence properties of the iterative optimization procedure, it can be shown that convergence is achieved when

$$w_{xi} = \frac{w}{\partial^2 U / \partial x_i}$$

(similarly for y_i) and $0 < w < 2$.

[0106] A fitting experiment has confirmed this.

[0107] The relaxation process described above behaves well and has been shown to approximate the surface very precisely. This is somewhat surprising, because the set of data points 95 is not used during the mesh relaxation. The basis for this surprising result lies in the highly constrained nature of a developable surface or its discrete approximation such as the mesh 110 with the contstraint that for each node i, $d_{ij} = const$. When the mesh is initialized onto the data, the nodes do not satify this constraint. However, the relaxation procedure causes nodes to be displaced orthogonally to satisfy the contraints. The form of the surface does not change drammatically, which would be the case if the displacements were tangent. This key observation is what makes the relaxing mesh approximate the surface without data.

[0108] Once the mesh is fitted properly to the three-dimensional data, the next phase is to texture-map the initial planar mesh. As we mentioned before, with this technique there is no need to unroll the surface just fitted, because we already have it to start with.

[0109] Texture-mapping to de-warp the curled document consists of three phases, which are illustrated in Figure 22.

First, all tiles 111 in the planar mesh 110 are initialized and relaxed 150 to the characterised document surface 152 such as to keep isometry. Using the known imaging geometry, the tile 111, which now lies on the three-dimensional surface 152, is back-projected 154 to the image plane 156 so as to obtain the texture 158 from the image that correspond to it. The final phase is to warp 160 the tile texture back 158 to its corresponding planar tile 111 so as to restore the texture as if it had been captured originally in a flat state.

[0110] The warp stage 160 is a standard process and a number of excellent algorithms are described by George Wolberg in an book titled *Digital Image Warping,* published by IEEE Computer Society Press, 1991.

[0111] The document curl characterisation approach described above works well regardless of the paper type and document complexity. The invention provides a practical and cheap means of characterising and de-warping page curl. In particular, the profile of the imaged document is determined by projecting a known two-dimensional structured light pattern and triangulating with the image of the pattern taken by the camera. The use of a two-dimensional pattern, as opposed to a single stripe or point, is particularly desirable in this applications because it does not require expensive moving parts (and their drivers) and allows quick characterisation of the page-curl in a single shot, not by sweeping a single beam over the page.

[0112] In this example, the stripes are identified only from images of the plain stripes, without moving either the detector array or the lights stripe projector or imposing any kind of temporal or spatial modulation, which would increase heavily on the system cost.

[0113] The method described above for characterising document curl in a platenless document scanning system is practical and fast and can be implemented with relatively inexpensive hardware for a document imaging application.

[0114] The document curl correction method presented above uses a mathematical model of paper, and an effective initialization and relaxation process for fitting this model to the data in a way that naturally produces an undistorted image. This is accomplished despite the fact that there are a large number of degrees of freedom and a large number of constraints that need to be satisfied with sparse and noisy data. This method has the ability to interpolate, extrapolate and self-complete wherever data is missing. The method produces high quality de-warped images of curled documents by modelling paper deformation in a physically realistic way.

Claims

1. An imaging system (1) for imaging a non-planar applicable surface (30), the system (1) comprising a processor (25) linked to an image capture means (2) and being capable of: capturing at least one image (31,33) of the surface (30), said image (31,33) having a warp corresponding to the non-planar surface; and of generating therefrom a first set of data points (95) representing the three-dimensional profile of the non-planar surface relative to a planar reference surface (12), wherein the processor .(25) is arranged to fit to the first set of data points (95) a second set of data points (117) representative of an applicable mesh (110) and to use the second set of data points (117) to texture-map the image (150,154,160) in order to de-warp the image (32).

2. A method of imaging a non-planar applicable surface (30) using an imaging system (1) comprising a processor (25) linked to an image capture means (2), comprising the steps of:

   i) capturing at least one image (31,33) of the surface (30) said image having a warp corresponding to the non-planar surface (30) ;

   ii) generating from the image (31,33) a first set of data points (95) representing the three-dimensional profile of the non-planar surface (30) relative to a planar reference surface (12);

   iii) fitting to the first set of data points a second set of data points (117) representative of an applicable mesh (110); and

   iv) using the second set of data points (117) to texture-map (150,154,160) the image (31,33) in order to de-warp the image.

3. A method as claimed in Claim 2, in which in step iii) the mesh (133) is distorted as the second set of data points (117) is fit to the first set of data points (95) to the extent that the mesh (133) is no longer applicable, following which the distorted mesh (133) is relaxed (140) to an applicable state.

4. A method as claimed in Claim 3, in which prior to step iii) an initial surface (130) is fit to the first set of data points, and in step iii) the mesh (133) is fit to the initial surface (130).

5. A method as claimed in Claim 4, in which the initial surface (130) is an applicable surface, and in which after fitting of the mesh (133) to the initial surface (130) at least some of the second set of data points (117) ar moved (132) closer to corresponding ones of the first set of data (95) points during which the mesh (133) is distorted.

6. A method as claimed in any of Claims 3 to 5, in which data points (134) in the second set of data points are not fit to the first set of data points (95) if said data points (134) in the second set of data points do not correspond closely enough to any of the data points in the first set of data points (95).

7. A method as claimed in any of Claims 3 to 6, in which the relaxation of the mesh (140) takes place in an iterative process in which the second set of data points (117) is adjusted incrementally until distances between points in the second set of data points are equalized.

8. A method as claimed in any of Claims 2 to 7, in which the non-planar applicable surface is a curled document (30).

9. A method as claimed in Claim 8, in which the extent of the document (30) is estimated by fitting a rectangle (124) around extreme points (126,128) of the first set of data points (95).

10. A method as claimed in Claim 8 or Claim 9, in which the image capture means (2) projects a structured light pattern (26) that forms separated light stripes (35) across the non-planar applicable surface (30), the first set of data points (95) being generated from the light stripes (35).

11. A method as claimed in Claim 10, in which step ii) includes the steps of:

a) creating a difference image by taking a difference between an image captured with the stripes (33) and an image captured without the stripes (31);

b) thresholding the difference image to discard portions below a threshold;

c) counting detected stripes across the difference image in order to identify individual stripes (35);

d) triangulating (84) the image of the non-planar surface at points (86) corresponding with identified stripes (35) to generate the first set of data points (95).

Fig. 1

Fig. 2

Fig. 3

Fig. 4

Fig 5
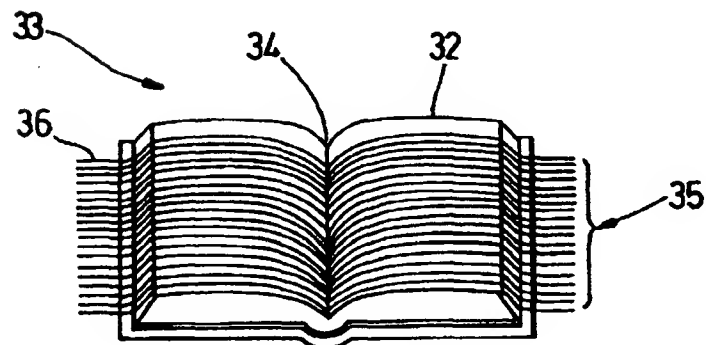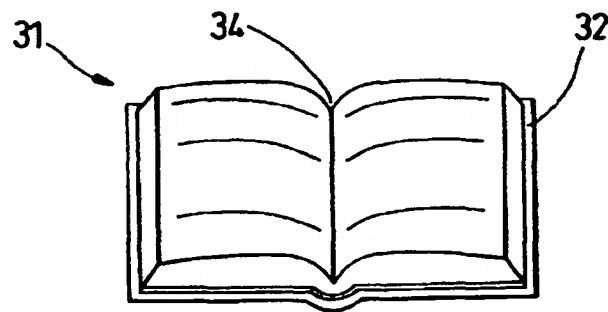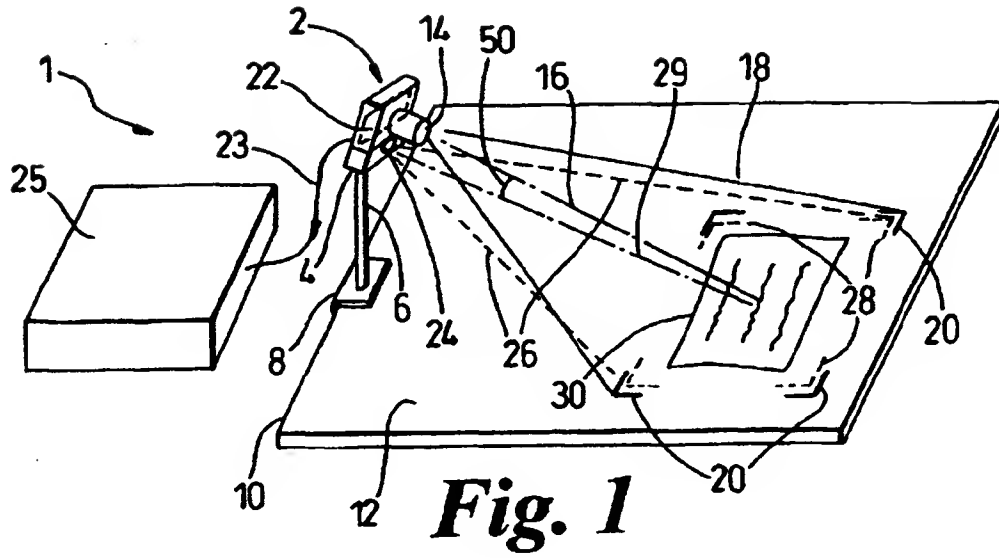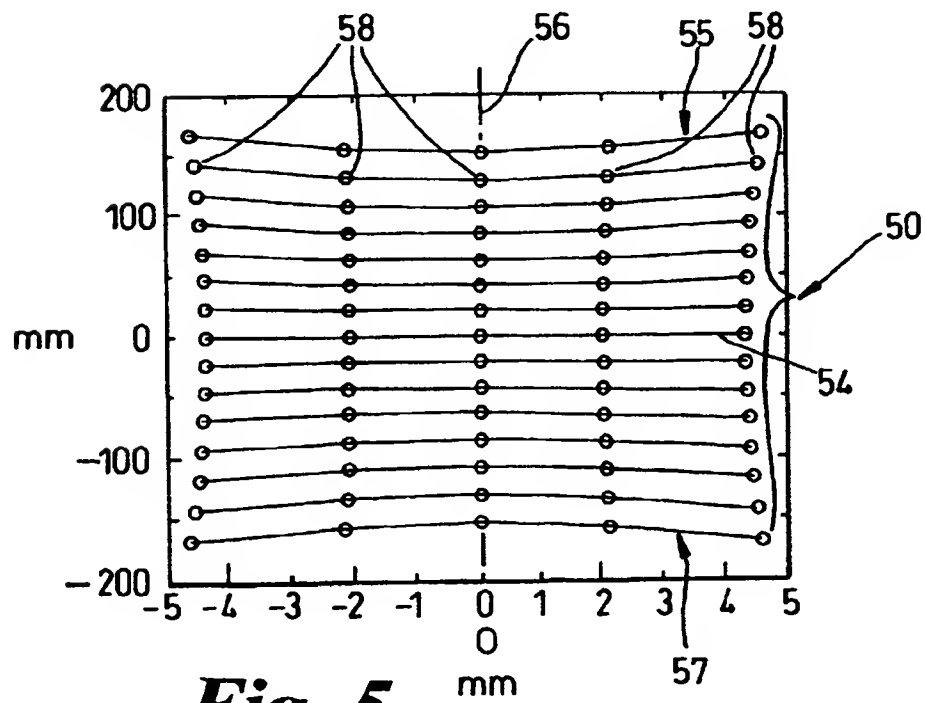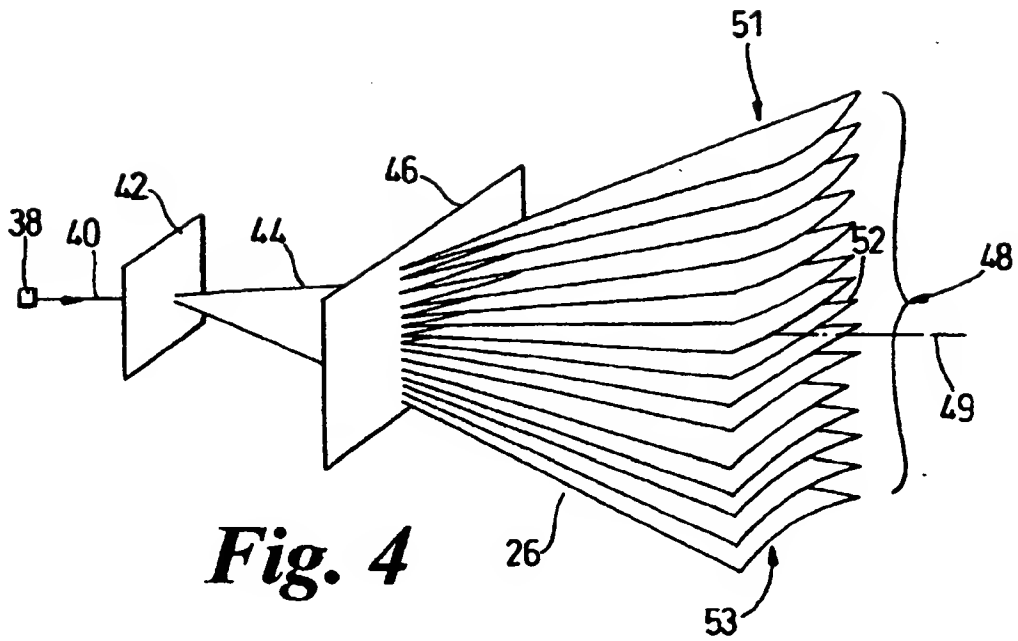
Fig. 6

Fig. 7

Fig. 8

Fig. 9A
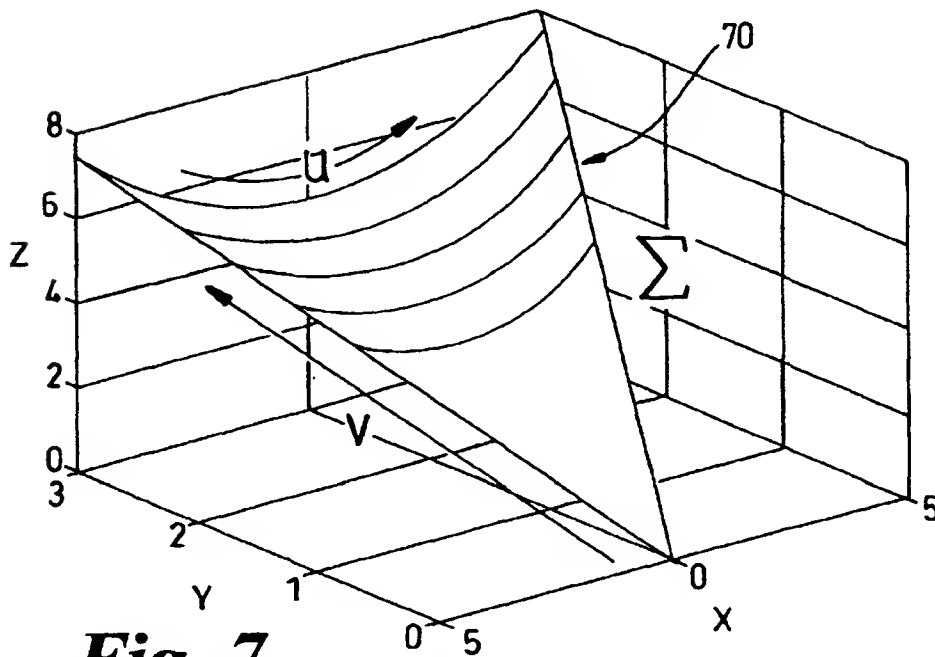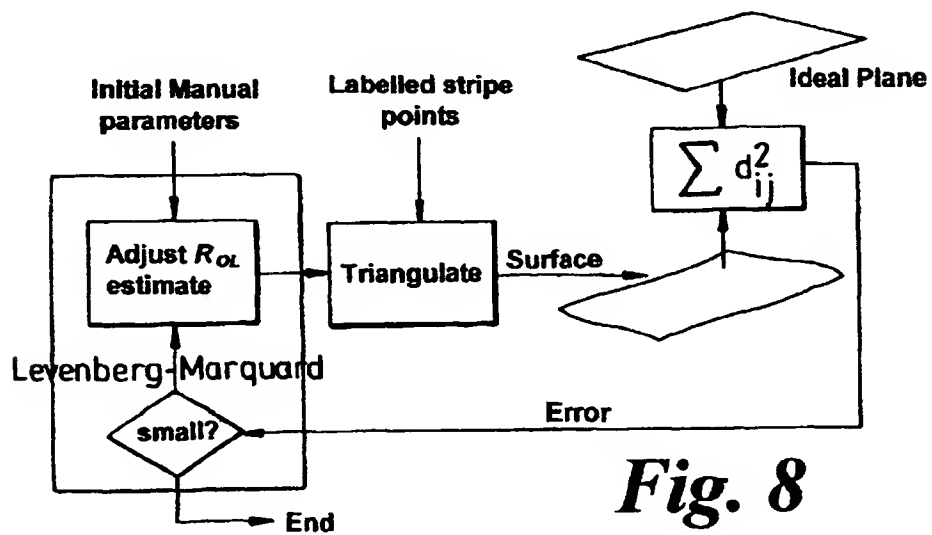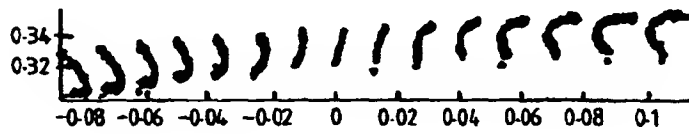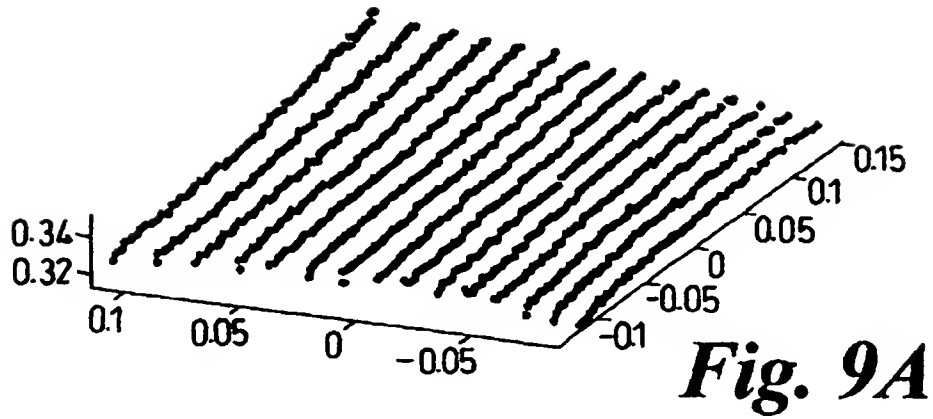
Fig. 9B
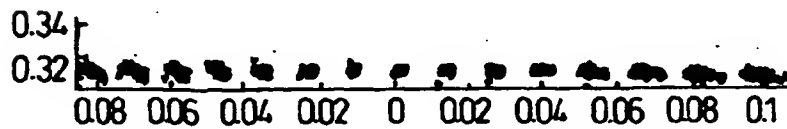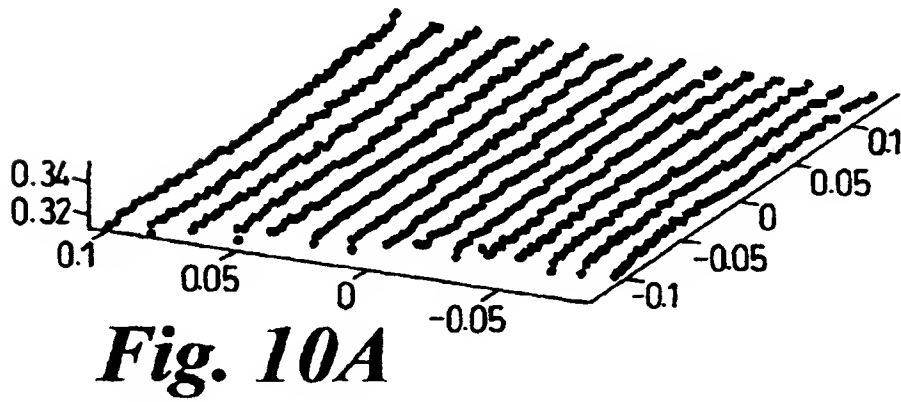
Fig. 10A

Fig. 10B
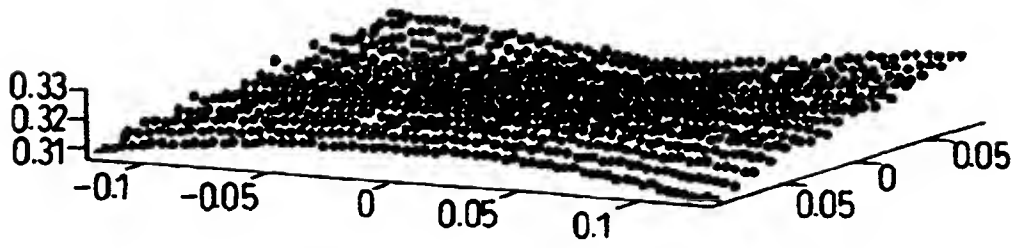
Fig. 11A

Fig. 11B

Fig. 12

$$^{o}P = (X, Y, Z)$$

$$^{o}p = (x_i, y_i)$$

$$R_{OL}$$
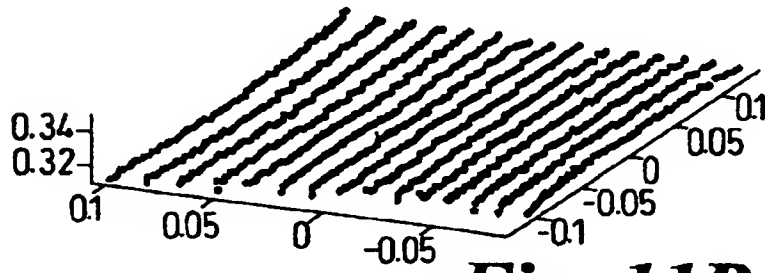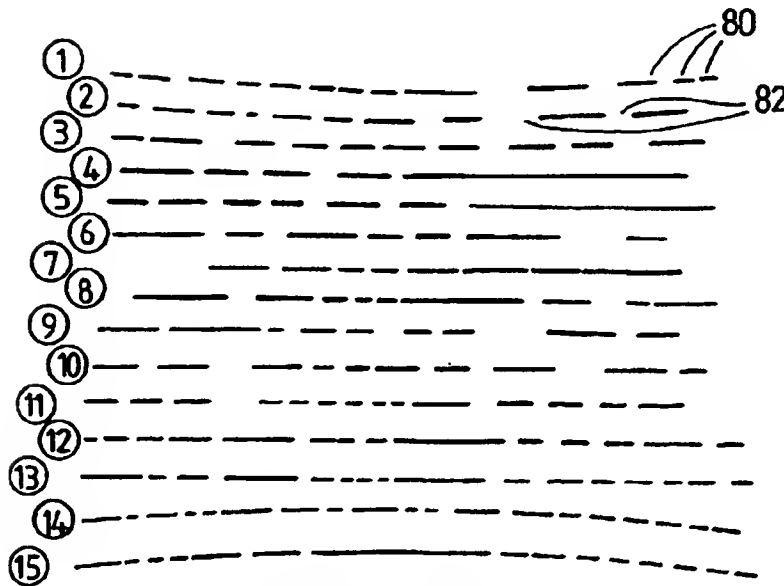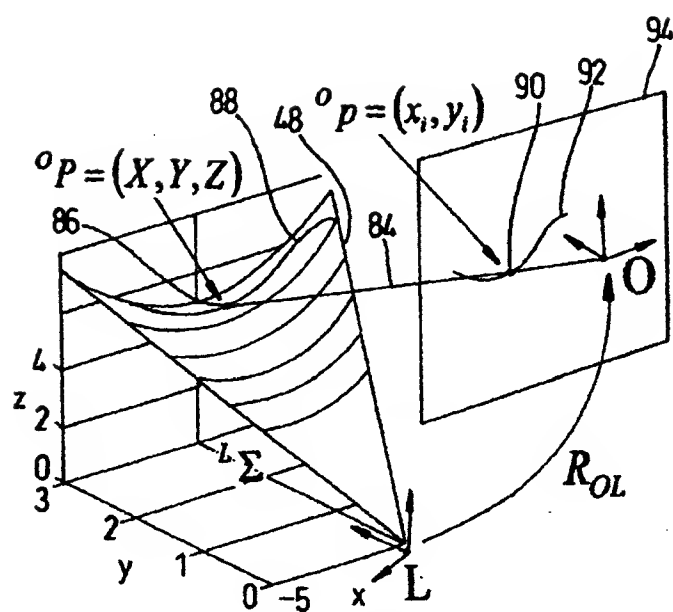
Fig. 13

Fig. 14

Fig. 15

Fig 16

Fig. 17

Fig. 18
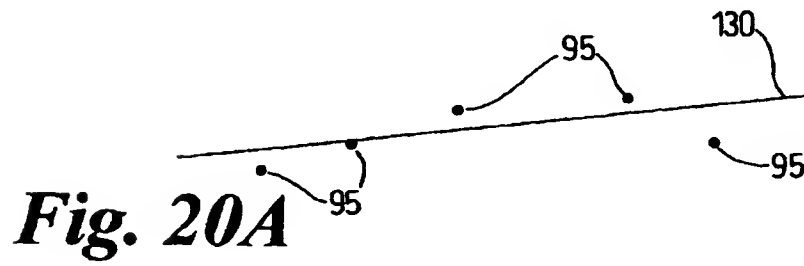
Fig. 19

Fig. 20A

Fig. 20B

Fig. 20C
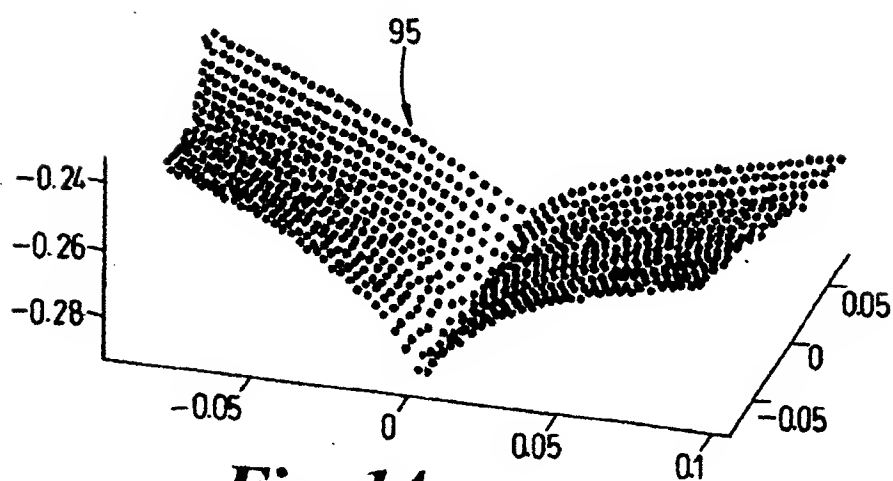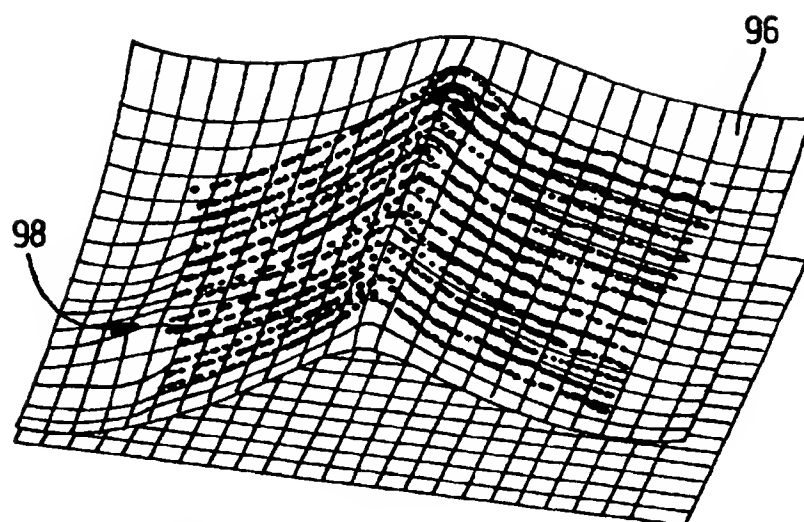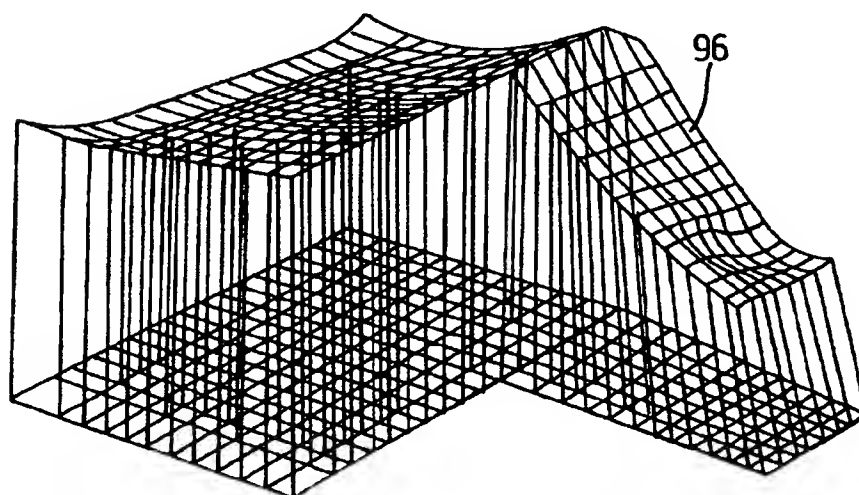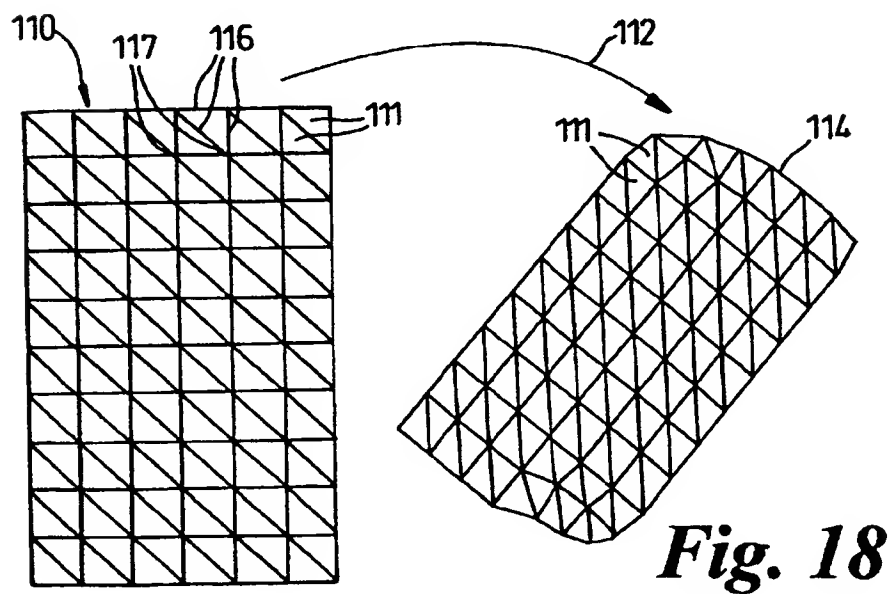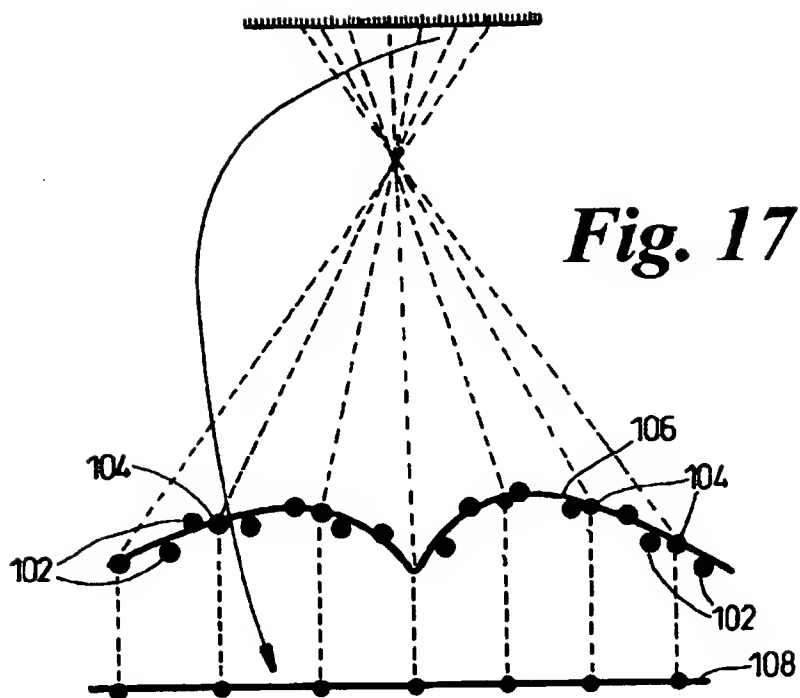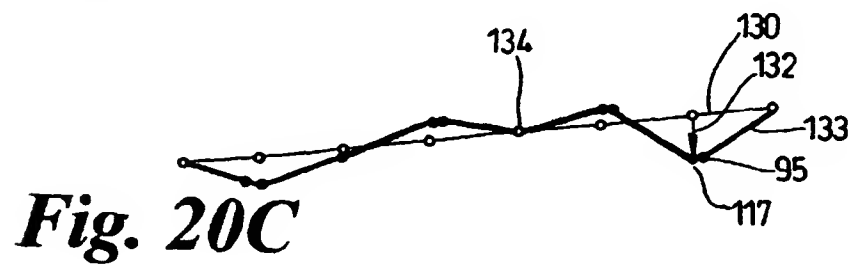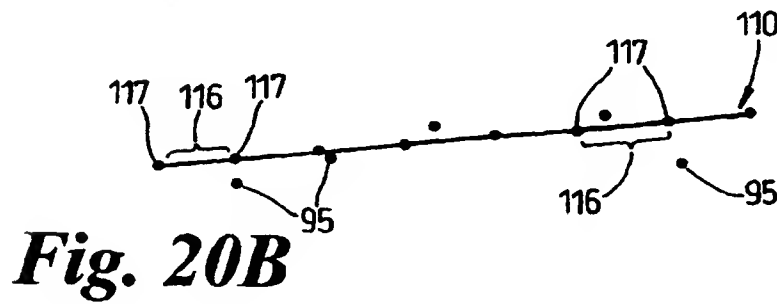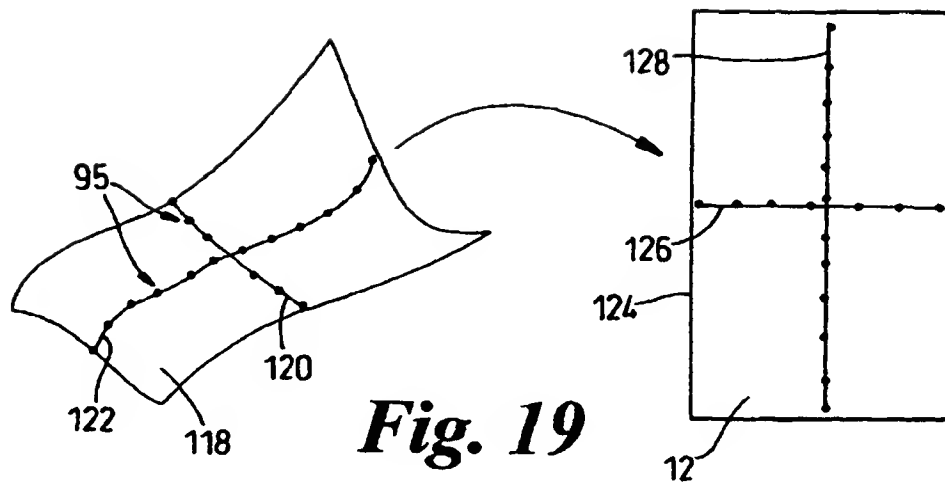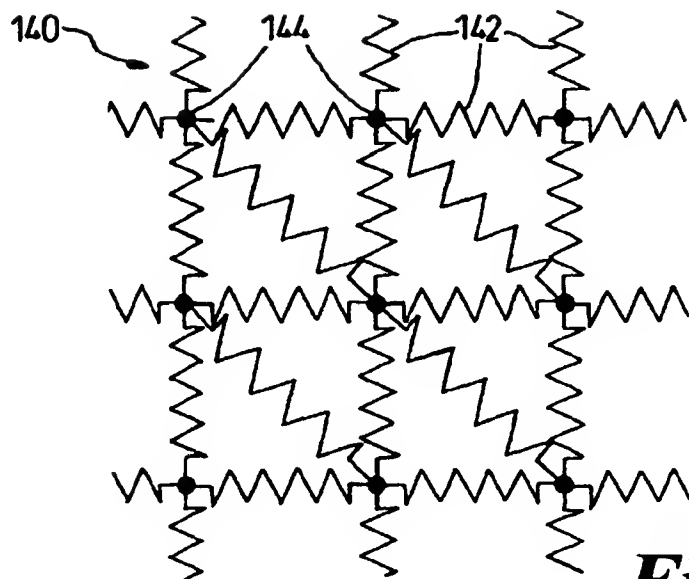
Fig. 21
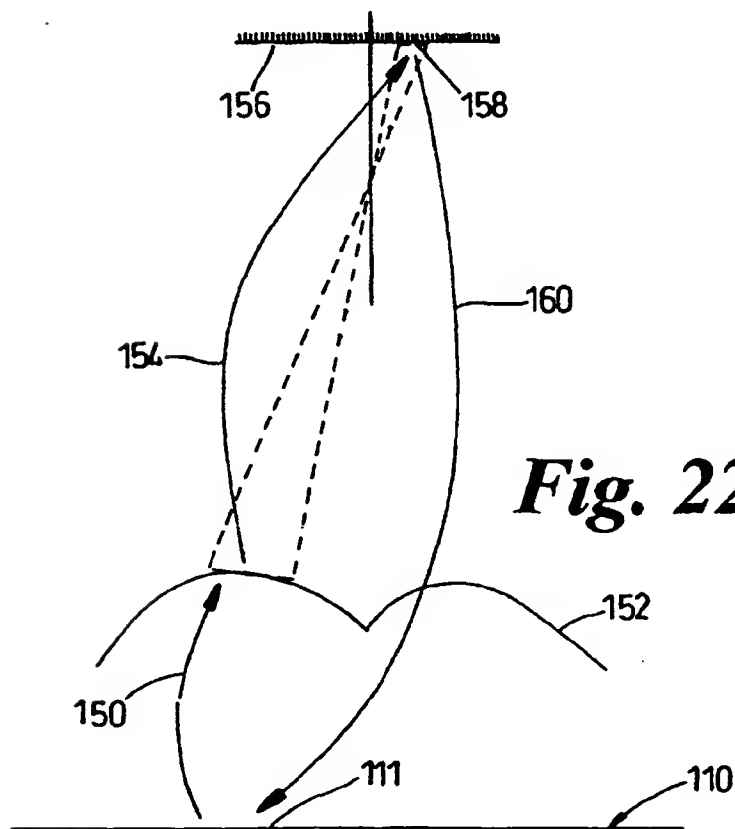
Fig. 22

European Patent Office

EUROPEAN SEARCH REPORT

Application Number

EP 99 30 5483

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
D,Y	US 5 760 925 A (BERLIN ANDREW A ET AL) 2 June 1998 (1998-06-02) * abstract; figures 1-13 * * column 6, line 19 - column 7, line 60 *	1,2	H04N1/047
Y	US 4 802 759 A (MATSUMOTO GORO ET AL) 7 February 1989 (1989-02-07) * abstract; figures 1-9 * * column 3, line 24 - column 6, line 54 *	1,2	
A	US 5 319 567 A (EBENSTEIN SAMUEL E) 7 June 1994 (1994-06-07) * column 11, line 1 - line 33 *	1,2	
A	US 5 585 926 A (FUJII SHINICHI ET AL) 17 December 1996 (1996-12-17) * abstract *	1,2,8	
A	WO 95 13520 A (BRITISH NUCLEAR FUELS PLC ;LIN LI (GB); STEEN WILLIAM MAXWELL (GB)) 18 May 1995 (1995-05-18) * abstract; figure 8 * * page 16, paragraph 3 - page 17, paragraph 1 *	1,2	TECHNICAL FIELDS SEARCHED (Int.Cl.7) G01B
A	US 5 497 236 A (WOLFF GREGORY J ET AL) 5 March 1996 (1996-03-05) * abstract; figures 4,5 *	1,2	
A	WO 91 15732 A (INTELLIGENT AUTOMATION SYSTEMS) 17 October 1991 (1991-10-17) * figure 2 *	10	

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
BERLIN	16 September 1999	Kassow, H

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.

EP 99 30 5483

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

16-09-1999

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
US 5760925	A	02-06-1998	NONE		
US 4802759	A	07-02-1989	JP	1892525 C	26-12-1994
			JP	6015968 B	02-03-1994
			JP	63044107 A	25-02-1988
US 5319567	A	07-06-1994	US	5384717 A	24-01-1995
			US	5414647 A	09-05-1995
			US	5442572 A	15-08-1995
			US	5481483 A	02-01-1996
US 5585926	A	17-12-1996	JP	5161000 A	25-06-1993
			JP	5161001 A	25-06-1993
			JP	5161002 A	25-06-1993
			JP	5161003 A	25-06-1993
			JP	5161004 A	25-06-1993
WO 9513520	A	18-05-1995	EP	0678189 A	25-10-1995
			JP	8505703 T	18-06-1996
			US	5714762 A	03-02-1998
US 5497236	A	05-03-1996	JP	7095387 A	07-04-1995
WO 9115732	A	17-10-1991	US	5589942 A	31-12-1996
			CA	2079817 A	06-10-1991
			DE	69127968 D	20-11-1997
			DE	69127968 T	12-02-1998
			EP	0523152 A	20-01-1993
			JP	5509394 T	22-12-1993